Australian Bureau of Statistics

1504.0 - Methodological News, Jun 2009
Previous ISSUE Released at 11:30 AM (CANBERRA TIME) 03/07/2009

Summary

Contents

Contents

New synthetic estimates enhance the retail time series
Update on the implementation of ANZSIC 2006 for sub-annual time series
Does ICT enable innovation?
Report on the Low Consumption Possibilities research project
Combining data from multiple surveys - LFS and NATSIHS
Combining data from multiple surveys - the SEW Data Pooling Project
Making Quality Visible Update
How to Contact Us and Subscriber Emailing List

New synthetic estimates enhance the retail time series

New synthetic estimates enhance the retail time series

No, it's not what you're thinking - sales of nylon, polyester and spandex have always been included in the Australian Retail time series. What is new is the synthetically estimated State-by-industry subgroup values for July - October 2008.

You may recall the article *Retail Survey Redesigned to Reduce Sample; Time Series Implications Considered* from the September 2008 edition of *Methodological News.* It outlined reductions to the sample size of the Retail Business Survey, the impacts on ABS Retail time series and the introduction of a quarterly Retail publication. While the impacts on the broad level monthly time series were small, fine State-by-industry subgroup level series exhibited strong survey design effects and increased volatility from July 2008. As a result, the release of state-by-industry series was restricted to the quarterly publication.

In November 2008, the Statistician re-instated the full sample size, citing increased scrutiny of economic statistics due to global financial developments and key users identifying more robust monthly retail data as their top priority. This meant that focus returned to the detailed monthly Retail estimates and the ceasing of the quarterly Retail publication after December 2008.

To ensure the consistency and continuity of the ABS monthly Retail time series, the Time Series Analysis section (TSA) re-estimated the fine State-by-industry subgroup level series for July to October 2008, which was a challenge because the small sample size and sample design effects made the series out of character for this period. The re-estimation was done by using historical State-by-industry subgroup data from July 1994 to June 2008 in conjunction with the November 2008 estimate. Treating the July - October estimates as missing, their values were interpolated using an application of the Kalman filter. These initial synthetic estimates then had an iterative reconciliation procedure applied to them to restore additivity up to the less affected higher aggregate marginal values. The final synthetic estimates were published in the March 2009 edition of *Retail Trade, Australia* (8501.0).

For more information, please contact Sam Allingham on (02) 6252 5718 or sam.allingham@abs.gov.au or Kirk Hampel on (02) 6252 5659 or kirk.hampel@abs.gov.au.

Update on the implementation of ANZSIC 2006 for sub-annual time series

Update on the implementation of ANZSIC 2006 for sub-annual time series

ANZSIC 2006 (A06) is an update of the current ANZSIC 1993 (A93) industry classification. This new classification will be applied to all sub-annual time series in the ABS from the September quarter 2009. The classification change will affect most ABS time series published on an industry basis. As A06 is a methodological change and is not related to any real world impact, A06 changes to the characteristics of ABS time series should be removed. In this way time series will remain consistent and comparable across time, and will be more useful for analysis and seasonal adjustment purposes.

Consistent A06 time series are being produced by re-estimating historical time series data collected under the old A93 industry classification, with the intention that they will be as close as possible to the new A06 classifications. Impact estimates to measure any remaining discrepancies between these base series and A06 estimates will be produced during a parallel process which will provide overlapping A06 and A93 estimates. Once the base series have been created and the impact estimates produced, the remaining discrepancy can be backcast out so that a single consistent and comparable A06 time series is produced.

To date, base series estimates have been created for all sub-annual collections and these new time series are currently undergoing seasonal reanalysis. Assessment of the backcast estimates will consider the amount of distortion to the seasonally adjusted movements in the base series, so backcasting should not be performed until seasonally adjusted estimates are available.

Backcasting will be run as a process to smooth out any remaining discrepancies between the new A06 estimates and the base series estimate (this difference is referred to as the A06 impact estimate). The Generalized Backcasting Facility (GBF) is a component of the SEASABS suite that can be used to run backcasting automatically across a group of time series, maintaining relevant aggregative relationships between these series and monitoring the effect of the backcasting process on the integrity of our estimates. Backcasting will be run initially using the impact estimate produced during the first parallel period and these first estimates will be used to test production systems and begin quality assurance of the backcast estimates.

The backcasting process will be iterated as updated impact estimates become available during the following parallel periods and if any updates need to be made to the base series estimates. The final backcasting run will be performed on September Quarter (SQ) estimates for a release of all sub-annual A06 estimates in SQ 2009 publications.

For more detail on the implementation of ANZSIC 2006 for sub-annual time series, please contact Nick von Sanden on (02) 6252 5727 or nick.von.sanden@abs.gov.au.

Does ICT enable innovation?

Does ICT enable innovation?

Innovation is a major driver of productivity gains in business, which facilitates not only improved performance for the individual firm but also continued, strong growth for the economy. Therefore, there is substantial interest in knowing what factors enable or drive business innovation in Australia. A research project to look into this is underway in the Analytical Services Branch (ASB).

The objective of the ASB study is to use the first two waves of data from the Business Longitudinal Database (BLD) to analyse the links between the use of information and communications technology (ICT) by businesses, and innovation. The use of the BLD provides an opportunity to look at the dynamics of the relationship between the two. Innovation and ICT use are also the subject of an OECD cross-country comparative study, to which this project will contribute its findings.

The primary question being addressed by this study, which is being undertaken with the help of the Innovation and Technology National Statistical Centre and the Economic Analysis and Reporting Branch, is whether ICT use is a source of business innovation, when controlling for other characteristics of the business. ICT use refers to things such as the use of broadband Internet, business web presence, employment of IT staff, business use of e-commerce, and automated system linkages. The project is testing the hypothesis that businesses which use ICT more intensively are more likely to innovate, and furthermore, are likely to innovate more intensively.

Specific types of innovation are being analysed (product, operational process, organisational/managerial and marketing, or the combinations of these). The analysis also considers the different stages of innovation (implemented, ongoing or abandoned), the degrees of novelty, and also according to whether the innovation was developed internally or in consultation with collaborative partners.

Following a review of the theoretical and empirical literature, exploratory analysis of the BLD was undertaken, in which an appropriate sample of businesses was selected. Initial descriptive analysis focused on the characteristics of those businesses which were innovating, and investigated businesses which became, or ceased to be, innovative between 2005-06 and 2006-07. Following detailed cross-tabulation and correlation analysis, the study is

now developing regression models to examine the links between the different measures of innovation and ICT use.

For more information, please contact Jessica Todhunter on (02) 6252 6992 or jessica.todhunter@abs.gov.au..

Report on the Low Consumption Possibilities research project

Report on the Low Consumption Possibilities research project

The Analytical Services Branch (ASB) and the Living Conditions Section have just completed the analytical phase of the "Low Consumption Possibilities" project. The main objective of the project is to develop a conceptual framework of low consumption possibilities based on income and wealth. The framework will inform the development of measures that will be used to more accurately identify Australian households that have low consumption possibilities and are at risk of experiencing economic hardship. “Consumption possibilities” refers to people’s command over resources that can be used to obtain goods and services to satisfy their needs and wants.

The use of low consumption possibilities to identify economic hardship is a conceptual shift away from previous efforts which used income alone as the basis for identifying economic hardship. This shift has been influenced, in part, by the fact that income is not a good predictor of living standard outcomes if it is measured by what people consume. ABS Income and Expenditure Survey data have shown that some households in the lowest income deciles quite often have expenditure that is higher than their reported income, while others have higher expenditures than households in the second or third income deciles. This probably suggests that some of the households in the bottom income deciles may not be experiencing economic hardship or that they may be financing their consumption from assets, debt or from some sort of transfers. This situation underlines the role of wealth as an important component of consumption and well-being, and the need to go beyond income alone as a determinant of economic well-being and economic hardship.

There are also practical and conceptual reasons for this shift. First, the term “consumption possibilities” more reliably reflects the economic resources available to a household than measures of income (or wealth) alone. Second, the concept of consumption possibilities considers the household’s capacity to consume, rather than just its actual or current consumption.

Although the role of wealth in consumption and well-being has long been noted, the absence of comprehensive data on wealth has precluded the use of wealth as a component in the measurement of economic hardship and well-being. This situation changed since data from the integrated 2003–04 *Survey of Income and Housing* (SIH) and *Household Expenditure Survey* (HES) became available. The HES collected data on the expenditure, income, net worth and other characteristics of persons resident in private dwellings throughout Australia. It included an expanded range of questions on income, and for the first time, it also included a comprehensive range of questions on household assets and liabilities.

Two methods were developed for the measurement of consumption possibilities. Both involved a consideration of the joint distribution of income and wealth. The first method, called the annuitisation method, involves the integration of the separate distributions of income and wealth into one distribution. The second method, called the ‘Low Economic Resources’ (LER) method, involves the identification of households that are located simultaneously in the bottom four deciles of both the income and wealth distributions. The annuitisation method yielded estimates of equivalised wealth-adjusted income (EWAI) for persons living in households. Persons living in households in the bottom two deciles of EWAI were deemed to have low economic resources. Modelling and analyses of the characteristics of persons at various points in the income and wealth distributions were used in selecting the cut-off points for both LER and the EWAI distributions.

The results of the analysis showed a very close overlap between the two methods in the types of persons identified as experiencing economic hardship. About 85 per cent of persons identified by the annuitisation method as having low wealth-adjusted income are also identified by the LER method as having low economic resources and at risk of economic hardship. On the whole, about 17 per cent of the population are identified as having both low EWAI and LER. On the other hand, nearly five per cent of the population are identified as having low EWAI but not LER, while 3.4 per cent are identified as having LER but not low EWAI. About 75 per cent of the population has neither low EWAI nor LER.

Validation of the methods showed that the characteristics of persons identified by both methods as having low consumption possibilities are consistent with characteristics normally associated with economic hardship. These persons are more likely than other persons to depend on government pensions and allowances as their principal source of current household income, to live in public housing, to not be in the labour force and to report high levels of financial stress and financial management problems. Demographically, higher than average proportions of persons identified by the two methods as having low consumption possibilities live in one-parent households with dependent and non-dependent children as well as in lone-person and couple-only households where the reference person is aged 65 years and over. They are also likely to have lower levels of educational attainment than persons identified as not having low consumption possibilities.

For more information, please contact Tetteh Dugbaza on (02) 6252 7221 or tetteh.dugbaza@abs.gov.au.

Combining data from multiple surveys- LFS and NATSIHS

Combining data from multiple surveys- LFS and NATSIHS

The ABS has been investigating the feasibility of combining data from multiple surveys to improve estimates of population totals. The key advantage in combining data is increased sample size, which reduces sampling error. However, inconsistencies between surveys, in areas such as scope, sample design and questions, may increase non-sampling error if they are not accounted for when combining data.

As reported in *Methodological News* two quarters ago, a research project in the Analytical Services Branch (ASB) is using the Labour Force Survey (LFS) and the National Aboriginal and Torres Strait Islander Health Survey (NATSIHS) to evaluate the benefits and issues involved in combining data from two surveys. The surveys were combined to produce labour force estimates for the Indigenous population in Australia. Both surveys collect labour force information but use a different set of questions to ascertain a respondent's labour force status. The LFS questions are more detailed.

Three approaches to combining the data were considered. In Approach 1, the LFS and NATSIHS labour force variables were assumed to be consistent and combined to produce a labour force estimate. In Approach 2 and Approach 3, the LFS labour force variable was taken as the 'gold-standard' and the NATSIHS labour force variable was assumed to contain some measurement error. In Approach 2, a 'NATSIHS' labour force variable was imputed for each LFS respondent, which allowed the use of a two phase estimator. In Approach 3, a 'LFS' labour force variable was imputed for each respondent to the NATSIHS, which was then combined with the LFS to produce a labour force estimate. In some cases the imputation was stochastic and respondents were assigned a probability of being employed, unemployed or not in the labour force.

All three approaches produced Australia-level labour force estimates with lower Relative Standard Errors (RSEs) than the corresponding LFS and NATSIHS estimates. Each approach also produced some substantial RSE gains for lower-level estimates, such as at the state by area (major city, regional or remote) level.

A set of diagnostics was developed to assess the quality of the estimates produced by combining the surveys and these were applied to Approach 1. One diagnostic showed that the employment and unemployment estimates at the state by area level in Approach 1 are more efficient than the corresponding LFS estimates, as long as the true difference between the LFS and NATSIHS estimates are less than 15% and 30% respectively. The conclusion of the work is that it is worthwhile to combine the LFS and NATSIHS**.** The diagnostics will soon be applied to Approach 2 and 3 to determine if they are more robust than the first.

For more information, please contact James Chipperfield on (02) 6252 7301 or james.chipperfield@abs.gov.au, or Julia Chessman on (02) 6252 5098 or julia.chessman@abs.gov.au

Combining data from multiple surveys- the SEW Data Pooling Project

Combining data from multiple surveys- the SEW Data Pooling Project

Users of official statistics are becoming more sophisticated, requiring estimates at more disaggregated or smaller subpopulation levels. However, most often individual surveys are not large or extensive enough to provide estimates at the desired level, and budgetary pressures preclude carrying out separate, smaller surveys for each analytical problem. There is potentially great value, therefore, in pooling or combining data from existing sources to construct new estimates at the required levels. In addition to the work described in the previous article, the Analytical Services Branch (ASB), in conjunction with the National Centre for Education and Training, has begun an investigation into the feasibility of pooling or combining data from different education surveys to derive better estimates of selected key educational participation and attainment measures.

The main survey of interest is the ABS Survey of Education and Work (SEW), which is being used by policy departments to produce key performance measures of Australian youth participation and attainment in education and training. There is strong interest in the disaggregation of these measures by state/territory and in movements over time (year to year). As a supplementary survey to the Labour Force Survey, SEW delivers accurate point-in-time estimates at the national level and reasonably accurate estimates for almost all states/territories, with the exception of Northern Territory. In addition to providing more accurate estimates at state/territory level and possibly at other relevant subpopulation grouping levels (e.g., sex, age, area of socioeconomic disadvantage), stakeholders are also interested in the ability of the data to detect relatively small movements in the key performance measures from year to year. The relatively small sample sizes for the required variables from the current SEW surveys do not

allow detection of small year-to-year movements, particularly for smaller jurisdictions and subpopulations.

The aim of the SEW Data Pooling project is therefore to assess the feasibility and benefits of pooling or combining SEW data with historical SEW surveys and/or other surveys, in order to improve the accuracy of the key performance measures of participation and attainment. Broadly, this project will investigate the improvements in accuracy (i.e., SEs/RSEs) of both single year and movement estimates of the key performance measures, under different options, relative to the accuracy achieved from estimates based only on SEW. That is, improvements will be recognised in comparison to accuracy achieved for estimates based only on SEW.

Initially, the following four key COAG educational measures will be the focus of analytical comparison: the proportion of 18-24 year olds engaged in [full time] employment, education or training at or above Certificate III level; the proportion of 19 year olds who have completed Year 12 or equivalent or Certificate II or above; the proportion of 20 to 24 year olds who have completed Year 12 or equivalent or Certificate II or above; and the proportion of 25 to 29 year olds who have completed Certificate III or above. These measures have been chosen due to their prominence in reporting against the COAG National Education Agreement, or the MCEETYA annual National Report on Schooling.

Possible different options for data pooling could include: combining current SEW with one or more previous SEWs; combining SEW with another monthly Labour Force supplementary survey (e.g., Labour Mobility Survey, Job Experience Survey, Underemployed Workers Survey, Childhood Education and Care Survey); combining SEW with another Special Social Survey (e.g., Adult Literacy and Life Skills, Survey of Education and Training, Survey of Disability, Ageing and Carers); and combining SEW with any combinations of the above three options.

If the initial phase proves the feasibility of the approach, then the second phase project will proceed with undertaking the actual pooling/combining of data to identify where improvements are possible, and make recommendations regarding the most suitable data pooling/combining process and option which will deliver the greatest accuracy improvements in the identified key educational performance measures.

For more information, please contact Anil Kumar on (02) 6252 5344 or anil.kumar@abs.gov.au

Making Quality Visible Update

Making Quality Visible Update

In the March 2009 edition of *Methodological News*, an update was provided on the Making Quality Visible initiative underway in the ABS. This article focuses on one aspect of the Making Quality Visible project, which is the Australian Bureau of Statistics Data Quality Framework (ABS DQF), recently released on the ABS website (ABS Data Quality Framework cat. no. 1520.0).

The ABS DQF is based on the Statistics Canada Quality Assurance Framework (2002) and the European Statistics Code of Practice (2005). It is comprised of seven dimensions of quality which are: Institutional Environment, Relevance, Timeliness, Accuracy, Coherence, Interpretability, and Accessibility.

The ABS DQF can be used to:

- define the quality of a data item or collection of data items;
- assess data in the context of a data need; and
- identify data gaps and areas for future improvement.

It is important to note that the level of importance assigned to each dimension of the ABS DQF is subjective and will depend on the purpose of the person using it.

The following gives an overview of the seven dimensions of quality:

Institutional Environment- refers to the institutional and organisational factors which may have a significant influence on the effectiveness and credibility of the agency producing the statistics.

Relevance- refers to how well the statistical product or release meets the needs of users in terms of the concept(s) measured, and the population(s) represented.

Timeliness- refers to the delay between the reference period (to which the data pertain) and the date at which the data become available; and the delay between the advertised date and the date at which the data become available (i.e., the actual release date).

Accuracy - refers to the degree to which the data correctly describe the phenomenon they were designed to

measure.

Coherence - refers to the internal consistency of a statistical collection, product or release, as well as its comparability with other sources of information, within a broad analytical framework and over time.

Interpretability - refers to the availability of information to help provide insight into the data.

Accessibility - refers to the ease of access to data by users, including the ease with which the existence of information can be ascertained, as well as the suitability of the form or medium through which information can be accessed.

For more detail on the dimensions and uses of the ABS DQF please see the ABS Data Quality Framework (cat. no. 1520.0), or contact . Narrisa GIlbert on (02) 6252 5283or narrisa.gilbert@ans.gov.au.

How to Contact Us and Subscriber Emailing List

How to Contact Us and Subscriber Emailing List

The Methodological Newsletter features articles and developments in relation to work done within the ABS Methodology and Data Management Division. By its nature, the work of the Division brings it into contact with virtually every other area of the ABS. Because of this, the newsletter is a way of letting all areas of the ABS know of some of the issues we are working on and help information flow. We hope the Methodological Newsletter is useful and we welcome comments.

If you would like to be placed on our electronic mailing list, please contact:

Jayne McQualter
Methodology & Data Management Division
Australian Bureau of Statistics
Locked Bag No. 10
BELCONNEN ACT 2617

Tel: (02) 6252 7320
Email: methodology@abs.gov.au

About this Release

The Methodological News is a quarterly information bulletin from the Methodology and Data Management Division. The newsletter features articles and developments in relation to work done with the division.

© Commonwealth of Australia

All data and other material produced by the Australian Bureau of Statistics (ABS) constitutes Commonwealth copyright administered by the ABS. The ABS reserves the right to set out the terms and conditions for the use of such material. Unless otherwise noted, all material on this website – except the ABS logo, the Commonwealth Coat of Arms, and any material protected by a trade mark – is licensed under a Creative Commons Attribution 2.5 Australia licence